NASA CR-
147826

(NASA-CR-147826) PHASE 4 OF THE RICE UNIVERSITY EARTH RESOURCES DATA ANALYSIS PROGRAM Final Report, Jun. 1975 - Jun. 1976 (Rice Univ.) 16 p HC $3.50 CSCL 05B

N76-27642

Unclas
G3/43 44568

ICSA

INSTITUTE FOR COMPUTER SERVICES AND APPLICATIONS

RICE UNIVERSITY

275-025-FPIV

Phase IV of the Rice University Earth Resources Data Analysis Program

FINAL REPORT

(June, 1975 - June, 1976)

Institute for Computer Services and Applications
Rice University
Houston, Texas 77001

This research sponsored by NASA under contract NAS 9-12776

I. Introduction:

During the past year, Rice University personnel have completed a variety of projects in two major areas. These areas are 1) the development of various pattern recognition algorithms and 2) the development of some general purpose algorithms useful in remote sensing calculations. In this report, we shall briefly describe the results of these projects and indicate the reports where more detailed information can be found. A list of all technical reports produced during the four years of this contract as well as abstracts of reports produced this year are contained in an appendix to this report.

II. Pattern Recognition Algorithm:

Three pattern recognition algorithms have been developed, implemented, and tested by Rice University personnel during the past year. These algorithms are used in 1) recursive estimation of mixture proportions, 2) optimal linear feature selection, and 3) estimation of probability densities.

In the first area, we have implemented and tested an algorithm proposed by D. Kazakos (ICSA report #275-025-019) of Rice University during Phase III of this contract. During this contract year, we have programmed this algorithm, with modifications as required, and tested it using both pseu-random data and Hill County LANDSAT data. The program was installed on the IBM 360/67 at LARS, Purdue, and a user's guide along with test results was produced (ICSA report #275-025-028). In this report, we discuss the need for scrambling the data whose proportions are to be estimated as well as the value to be used for the parameter L . Various effects of the algorithm are also discussed. On the test cases run, the algorithm yielded results comparable to some of the better proportion estimators being tested at NASA [1].

In the optimal linear feature extraction area, three reports were written (ICSA reports #275-025-022, -025, and -026). Two algorithms were developed and tested to select the single best linear feature for Gaussian pattern classes and the best m features. The optimal n-to-one transformation algorithm was tested on C1 flight line data, Hill County LANDSAT data, and pseudo-random data. In most cases, it was found to yield roughly equal (maximum likelihood) classification performance to that achieved by using the best two channels as selected by the Bhattacharyya distance criterion. The optimal m dimensional feature extraction algorithm yielded similar results. Additionally, numerical results were presented of cases where the optimal n-to-one transformation yielded superior classification performance compared to using all channels, due to small sample sizes.

Algorithms were developed in the area of probability density estimation. These included investigations into Parzen-type estimators when sample sizes are small, and the development, implementation, and testing of a discrete maximum likelihood penalized estimator. Details are formed in ICSA report #275-025-023. Additionally, a program and user's guide were produced and given to the EOD.

III. General Purpose Algorithms:

A variety of general purpose algorithms useful in remote sensing calculations were developed during the contract year. Several of these were described in various ICSA technical reports, while others were reported in user's guides (along with the programs) distributed to the EOD. In this area, particular attention was paid to both numerical stability and efficiency.

One area of investigation was constrained optimization for probability density function (PDF)-type constraints (i.e. $x_i \geq 0$, $\sum_{i=1}^{n} x_i = 1$). For the quadratic programming case (i.e. $\min_x \| Ax - b \|^2$), the approach of Lawson and Hanson [2] appears suitable. Programs utilizing their approach with these constraints were developed and a user's guide, along with the programs, were given to the EOD. For the more general case (i.e. $\min_x f(x)$) another algorithm was developed, programmed, and tested (see ICSA report #275-025-029). The algorithm is quite general and also computationally efficient and numerically stable. Programs and a user's guide were provided.

A report was written detailing three methods for computing and updating the modified Cholesky decomposition of a covariance matrix (see report #275-025-024) Programs for each of the methods were included. It was found that one of the methods (the composite-t method) offered exceptional numerical stability and an updating capability compared to other more commonly used methods.

In the area of computational complexity, some work was done on the applicability of the Winograd [3] matrix multiplication algorithm to various calculations in remote sensing (see report #275-025-027). The regions where a computational savings could be effected were described and their dependence on the ratio of the computer multiply to add time investigated. In this same area, a memorandum was issued concerning timing considerations for using the modified Cholesky decomposition in maximum-likelihood classification.

Lastly, a clustering algorithm developed by Zahn [4] was tested. The program and documentation were delivered to the EOD.

APPENDIX A

Rice University

Institute for Computer Services & Applications

Technical Reports

Produced under NASA Contract NAS 9-12776

Number	Title	Author(s)	Date
001	"A Mathematical Model Concerning Reflectance from a Row Crop"	R. K. Jaggi	Sept., 1972
002	"Spline Smoothing of Histograms by Linear Programming"	J. O. Bennett	Sept., 1972
004	"Power Spectral Density Estimation by Spline Smoothing in the Frequency Domain"	R. J. P. de Figueiredo and J. R. Thompson	Sept., 1972
005	"A Geometrical Interpretation of the 2n-th Central Difference"	R. A. Tapia	Sept., 1972
006	"Data Smoothing and Error Detection Based on Linear Interpolation"	V. M. Guerra and R. A. Tapia	Sept., 1972
007	"A Random Number Generator for Continuous Random Variables"	V. M. Guerra, R. A. Tapia, and J. R. Thompson	Sept., 1972
008	"The Use of the Modified Cholesky Decomposition in Divergence and Classification Calculations"	D. L. Van Rooy, M. S. Lynn, and C. H. Snyder	May, 1973
009	"Use of Spatial Information in Clas[illegible]ication of Remotely Sensed Data"	D. L. Van Rooy and M. S. Lynn	May, 1973
010	"Optimal Feature Extraction - the Two Class Case"	W. S. Hsia and R. J. P. de Figueiredo	May, 1973

Number	Title	Author(s)	Date
011	"Estimation of Multivariate Probability Density Functions Using B-Splines"	J. O. Bennett	May, 1974
012	"Error Detection & Data Smoothing Based on Local Procedures"	V. M. Guerra	May, 1974
013	"Optimal Design of an Unsupervised Adaptive Classifier with Unknown Priors"	D. Kazakos	May, 1974
014	"Optimal Linear and Nonlinear Feature Extraction Based on the Minimization of the Increased Risk of Misclassification"	R. J. P. de Figueiredo	June, 1974
015	"Classification by Means of B-Spline Potential Functions with Applications to Remote Sensing"	J. O. Bennett, R. J. P. de Figueiredo, and J. R. Thompson	May, 1974
016	"Nonparametric Maximum Likelihood Estimation of Probability Densities by Penalty Function"	G. F. de Montricher, R. A. Tapia, and J. R. Thompson	August, 1974
018	"An Analysis of Applications Development Systems for Remotely Sensed, Multispectral Data for the Earth Observations Division of the NASA Lyndon B. Johnson Space Center"	D. L. Van Rooy, R. M. Smith, and M. S. Lynn	Feb., 1974
019	"Recursive Estimation of Prior Probabilities Using the Mixture Approach"	Demetrios Kazakos	Sept., 1974
020	"The Resource Utilization Monitoring System for the Large Area Crop Inventory Experiment— a Recommended Approach"	Richard A. Schafer, David L. Van Rooy, M. Stuart Lynn	Feb., 1975

Number	Title	Author(s)	Date
021	"Data Base Design and Maintenance for the Resource Utilization Monitoring System for the Large Area Crop Inventory Experiment—a Recommended Approach"	R. A. Schafer	May, 1975
022	"An Algorithm for Optimal Single Linear Feature Extraction from Several Gaussian Pattern Classes"	S. A. Starks, R. J. P. de Figueiredo, and D. L. Van Rooy	November, '75
023	"Nonparametric Probability Density Estimation by Optimization Theoretic Techniques"	David Warren Scott	April, 1976
024	"On the Computation and Updating of the Modified Cholesky Decomposition of a Covariance Matrix"	D. L. Van Rooy	May, 1976
025	"Classification Improvement by Optimal Dimensionality Reduction When Training Sets Are of Small Size"	S. A. Starks, R. J. P. de Figueiredo, and D. L. Van Rooy	April, 1976
026	"An Algorithm for Extraction of More Than One Optimal Linear Feature from Several Gaussian Pattern Classes"	R. J. P. de Figuieredo, K. C. Pau, A. D. Sagar, S. A. Starks, and D. L. Van Rooy	April, 1976
027	"The Use of the Winograd Matrix Multiplication Algorithm in Digital Multispectral Processing"	D. L. Van Rooy, R. A. Schafer, and M. S. Lynn	May, 1976
028	"The Recursive Maximum Likelihood Proportion Estimator—User's Guide and Test Results"	D. L. Van Rooy	June, 1976
029	A Quasi-Newton Approach to Optimization Problems with Probability Density Constraints"	R. A. Tapia and D. L. Van Rooy	June, 1976

8

Number	Title	Author(s)	Date
275-025-FPI	Phase I of the Earth Resources Data Analysis Program--FINAL REPORT		June '72-May '73
275-025-FPII	Phase II of the Rice University Earth Resources Data Analysis Program--FINAL REPORT		June '73-May '74
275-025-FPIII	Phase III of the Rice University Earth Resources Data Analysis Program--FINAL REPORT		June '74-May '75
275-025-FPIV	Phase IV of the Rice University Earth Resources Data Analysis Program--FINAL REPORT		June '75-June '76

APPENDIX B

Abstracts of Rice University Institute for Computer Services & Applications Technical Reports Produced under NASA Contract NAS 9-12776 During Phase IV, 1976

"An Algorithm for Optimal Single Linear Feature Extraction from Several Gaussian Pattern Classes," #275-025-022, by S.A. Starks, R.J.P. de Figueiredo, and D.L. Van Rooy.

A computational algorithm is presented for the extraction of an optimal single linear feature from several Gaussian pattern classes. The algorithm minimizes the increase in the probability of misclassification in the transformed (feature) space. The general approach used in this procedure was developed in a recent paper by R.J.P. de Figueiredo [1].

Numerical results on the application of this procedure to the remotely sensed data from the Purdue C1 flight line as well as LANDSAT data are presented. It was found that classification using the optimal single linear feature yielded a value for the probability of misclassification on the order of 30% less than that obtained by using the best single untransformed feature. Also, the optimal single linear feature gave performance results comparable to those obtained by using the two features which maximized the average divergence.

"Nonparametric Probability Density Estimation by Optimization Theoretic Techniques," #275-025-023, by David Warren Scott.

In this study, two nonparametric probability density estimators are considered. The first is the kernel estimator. The problem of choosing the kernel scaling factor based solely on a random sample is addressed. An interactive mode is discussed and an algorithm proposed to choose the scaling factor automatically. In a Monte Carlo simulation study, the resulting integrated mean square error compares favorably with the error using the usual asymptotically optimal choice of the kernel scaling factor. For the latter case, the true sampling density is required to calculate the optimal scaling factor.

The second nonparametric probability estimate uses penalty function techniques with the maximum likelihood criterion. A discrete maximum penalized likelihood estimator is proposed and is shown to be consistent in the mean square error. Approximation results of this discrete solution to the corresponding infinite-dimensional solution are proved. A numerical implementation technique for the discrete solution is discussed and examples displayed. An extensive simulation study compares the integrated mean square error of the discrete and kernel estimators. The robustness of the discrete estimator is demonstrated graphically.

"On the Computation and Updating of the Modified Cholesky Decomposition of a Covariance Matrix," #275-025-024, by D.L. Van Rooy.

In this paper, we discuss three known methods for obtaining and updating the modified Cholesky decomposition (MCD) for the particular case of a covariance matrix when one is given only the original data. These methods are the standard method of forming the covariance matrix K then solving for the MCD, L & D (where $K = LDL^T$); a method based on Householder reflections; and lastly, a method employing the composite-t algorithm developed by Fletcher and Powell (Math Comp., 28, 1974, pp. 1067-1087). For many cases in the analysis of remotely sensed data, the composite-t method is the superior method despite the fact that it is the slowest one, since (1) the relative amount of time computing MCD's is often quite small, (2) the stability properties of it are the best of the three, and (3) it affords an efficient and numerically stable procedure for updating the MCD. The properties of these methods are discussed and FORTRAN programs implementing these algorithms are listed in an appendix.

"Classification Improvement by Optimal Dimensionality Reduction When Training Sets Are of Small Size," #275-025-025, by S.A. Starks, R.J.P. de Figueiredo, and D.L. Van Rooy.

When the sizes of the training sets are small, classification in a subspace of the original data space may give rise to a smaller probability of error than the classification in the data space itself. This is because the gain in the accuracy of estimation of the likelihood functions used in classification in the lower dimensional space (subspace) offsets the loss of information associated with dimensionality reduction (feature extraction). To test this conjecture, a computer simulation was performed. A number of pseudo-random training and data vectors were generated from two four-dimensional Gaussian classes. An algorithm previously described (ICSA Technical Report #275-025-022, EE Technical Report #7520) was used to create an optimal one-dimensional feature space on which to project the data. When the sizes of the training sets were small, classification of the data in the optimal one-dimensional space was found to yield lower error rates than the one in the original four-dimensional space. Specifically, depending on the sizes of the training sets, the improvement ranged from 11% to 1%.

"An Algorithm for Extraction of More Than One Optimal Linear Feature from Several Gaussian Pattern Classes," #275-025-026, by R.J.P. de Figueiredo, K.C. Pau, A.D. Sagar, S.A. Starks, and D.L. Van Rooy.

Two algorithms have been developed at Rice University for optimal linear feature extraction based on the minimization of the risk (probability) of misclassification under the assumption that the class conditional probability density functions are Gaussian. One of these algorithms, which applies to the case in which the dimensionality of the feature space (reduced space) is unity, has been described elsewhere [Rice University ICSA Technical Reports Nos. 275-025-022 and 275-025-025 (EE Technical Reports Nos. 7520 and 7603)]. In the present report, we describe the second algorithm which is used when the dimension of the feature space is greater than one.

Numerical results obtained from the application of the present algorithm to remotely sensed data from the Purdue C1 flight line are mentioned. Brief comparisons are made of these results with those obtained using a feature selection technique based on maximizing the Bhattacharyya distance. For the example considered, a significant improvement in classification is obtained by the present technique.

"The Use of the Winograd Matrix Multiplication Algorithm in Digital Multispectral Processing," #275-025-027, by D.L. Van Rooy, R.A. Schafer, and M.S. Lynn.

The Winograd procedure for matrix multiplication [S. Winograd, Comm. on Pure and Applied Math., 23, 1970] provides a method whereby general matrix products may be computed more efficiently than the normal method. In this report, we describe the algorithm and the time savings that can be effected. A FORTRAN program is provided which performs a general matrix multiply according to this algorithm.

Additionally, we describe a variation of this procedure that may be used to calculate Gaussian probability density functions. It is shown how a time savings can be effected in this calculation. The extension of this method to other similar calculations should yield similar savings.

"The Recursive Maximum Likelihood Proportion Estimator—User's Guide and Test Results," #275-025-028, by D.L. Van Rooy.

In this report, we describe our implementation of the recursive maximum likelihood proportion estimator proposed by D. Kazakos in "Recursive Estimation of Prior Probabilities Using the Mixture Approach," (Rice University, ICSA Technical Report #275-025-019). A user's guide to the programs as they currently exist on the IBM 360/67 at LARS, Purdue is included, and test results on LANDSAT data are described. On Hill County data, the algorithm yields results comparable to the standard maximum likelihood proportion estimator.

"A Quasi-Newton Approach to Optimization Problems with Probability Density Constraints," #275-025-029, by R.A. Tapia and D.L. Van Rooy.

A quasi-Newton method is presented for minimizing a nonlinear functional while constraining the variables to be nonnegative and sum to one. The nonnegativity constraints are eliminated by working with the squares of the variables and the resulting problem is solved using Tapia's general theory of quasi-Newton methods for constrained optimization. A user's guide for a program implementing this algorithm is provided.

REFERENCES

[1] W. A. Coberly and P. L. Odell, "An Empirical Comparison of Five Proportion Estimators," from the annual report of the University of Texas at Dallas for contract NAS 9-13512, January, 1975.

[2] C. L. Lawson and R. J. Hanson, Solving Least Squares Problems, Prentice-Hall, New Jersey, 1974.

[3] S. Winograd, "On the Number of Multiplications Necessary to Compute Certain Functions," Comm. on Pure and Applied Math., 23, 1970, pp. 165-179.

[4] C. T. Zahn, "Graph Theoretical Methods for Detecting and Describing Gestalt Clusters," IEEE Transactions on Computers, C-20, January, 1971, pp. 68-86.